Chain Rules for Smooth Min- and Max-Entropies

Alexander Vitanov, Frédéric Dupuis, Marco Tomamichel, and Renato Renner

***Abstract*—The chain rule for the Shannon and von Neumann entropy, which relates the total entropy of a system to the entropies of its parts, is of central importance to information theory. Here we consider the chain rule for the more general smooth min- and max-entropy, used in one-shot information theory. For these entropy measures, the chain rule no longer holds as an equality, but manifests itself as a set of inequalities that reduce to the chain rule for the von Neumann entropy in the i.i.d. case.**

***Index Terms*—smooth min- and max-entropies, chain rules.**

I. Introduction

In classical and quantum information theory, entropy measures are often used to characterize fundamental information processing tasks. For example, in his groundbreaking work on information and communication theory [1], Shannon showed that entropies can be used to quantify the memory needed to store the (compressed) output of an information source or the capacity of a communication channel. It follows immediately from the basic properties of the Shannon entropy that the equality

$$H(AB) = H(A|B) + H(B) \ ,$$

which we call the *chain rule*, must hold. Here, $H(B)$ denotes the entropy of the random variable B and $H(A|B)$ is the entropy of the random variable A averaged over *side information* in B. The chain rule therefore asserts that the entropy of two (possibly correlated) random variables, A and B, can be decomposed into the entropy of B alone plus the entropy of A conditioned on knowing B. More generally, one may average over additional side information, C, in which case the chain rule takes the more general form

$$H(AB|C) = H(A|BC) + H(B|C) \ . \tag{1}$$

The chain rule forms an integral part of the entropy calculus. The other basic ingredient is strong sub-additivity, which can be written as $H(A|BC) \leq H(A|C)$, i.e. additional side information can only decrease the entropy.

The quantum generalization of Shannon's entropy, the *von Neumann entropy*, inherits these fundamental properties. For a quantum state[1] ρ_A on A, the von Neumann entropy is defined as $H(A)_\rho := -\operatorname{tr}(\rho_A \log \rho_A)$, where tr denotes the trace and $\log$ is taken in base 2 throughout this paper. The conditional von Neumann entropy with classical side information can again be defined by an average, however, this intuitive definition fails if the side information is quantum. Pointing to its fundamental importance, the conditional von Neumann entropy is thus defined by the chain rule itself, i.e. $H(A|B)_\rho := H(AB)_\rho - H(B)_\rho$. In addition to the chain rule and strong sub-additivity, it also satisfies a duality relation: For any pure tripartite state ρ_{ABC}, we have $H(A|B)_\rho = -H(A|C)_\rho$.

Shannon and von Neumann entropies have been successfully employed to characterize an enormous variety of information theoretic tasks, many of which are of high practical relevance (examples include the aforementioned tasks of data compression or channel coding). However, a basic assumption usually made in this context is that the underlying random processes (e.g., those relevant for the generation of data, or the occurrence of noise in a communication channel) are modeled asymptotically by an arbitrarily long sequence of random variables that are *independent and identically distributed (i.i.d.)*. In the absence of this assumption (e.g., if a channel is only invoked a small number of times or if its noise model is not i.i.d.), the use of the von Neumann entropy is generally no longer justified. The formalism of smooth min- and max-entropy, introduced in [2]–[4] and further developed in [5]–[8], overcomes this limitation and enables the analysis of general situations beyond the i.i.d. scenario. This level of generality turned out to be crucial in various areas, e.g., in physics (where entropies are employed for the analysis of problems in thermodynamics [9]) or in cryptography (where entropies are used to quantify an adversary's uncertainty).

Smooth min- and max-entropy, denoted $H_{\min}^{\varepsilon}$ and $H_{\max}^{\varepsilon}$, respectively, depend on a positive real value ε, called *smoothing parameter* ε (see Section II for formal definitions). When the entropies are used to characterize operational tasks, the smoothing parameter determines the desired accuracy. For example, the smooth min-entropy, $H_{\min}^{\varepsilon}(A|B)$, characterizes the number of fully mixed qubits, independent (i.e. *decoupled*) from side information B, that can be extracted from a quantum source A [10], [11]. Furthermore, the smooth max-entropy, $H_{\max}^{\varepsilon}(A|B)$, characterizes the amount of entanglement needed between two parties, A and B, to merge a state ρ_{AB}, where ρ_A is initially held by A, to B [11], [12]. In both cases, the smoothing parameter ε corresponds to the maximum distance between the desired final state and the one that can be achieved.

Smooth entropy can be seen as strict generalization of Shannon or von Neumann entropy. In particular, the latter can be recovered by evaluating the smooth min- or max-entropy for i.i.d. states [3], [6]. Accordingly, smooth entropy inherits many of the basic features of von Neumann entropy, such as strong sub-additivity. In light of this, it should not come as a surprise that smooth entropy also obeys inequalities that generalize the chain rule (1). Deriving these is the main aim of this work.

F. Dupuis, R. Renner, and A. Vitanov are with the Institute for Theoretical Physics, ETH Zürich, 8093 Zürich, Switzerland. M. Tomamichel is with the Center for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117543. (e-mail: dupuis@phys.ethz.ch; renner@phys.ethz.ch; avitanov@student.ethz.ch; cqtmarco@nus.edu.sg)

[1]Formal definitions follow in Section II.

Specifically, one can obtain four pairs of generalized chain inequalities. For any small smoothing parameters $\varepsilon', \varepsilon'', \varepsilon''' \geq 0$ and $\varepsilon > \varepsilon' + 2\varepsilon''$, we have

$$H_{\min}^{\varepsilon}(AB|C)_\rho \geq H_{\min}^{\varepsilon''}(A|BC)_\rho + H_{\min}^{\varepsilon'}(B|C)_\rho - f\,,$$
$$H_{\max}^{\varepsilon}(AB|C)_\rho \leq H_{\max}^{\varepsilon'}(A|BC)_\rho + H_{\max}^{\varepsilon''}(B|C)_\rho + f\,,$$

$$H_{\min}^{\varepsilon'}(AB|C)_\rho \leq H_{\min}^{\varepsilon}(A|BC)_\rho + H_{\max}^{\varepsilon''}(B|C)_\rho + 2f\,,$$
$$H_{\max}^{\varepsilon'}(AB|C)_\rho \geq H_{\min}^{\varepsilon''}(A|BC)_\rho + H_{\max}^{\varepsilon}(B|C)_\rho - 2f\,,$$

$$H_{\min}^{\varepsilon'}(AB|C)_\rho \leq H_{\max}^{\varepsilon''}(A|BC)_\rho + H_{\min}^{\varepsilon}(B|C)_\rho + 3f\,,$$
$$H_{\max}^{\varepsilon'}(AB|C)_\rho \geq H_{\max}^{\varepsilon}(A|BC)_\rho + H_{\min}^{\varepsilon''}(B|C)_\rho - 3f\,,$$

$$H_{\min}^{\varepsilon'}(AB|C)_\rho \leq H_{\max}^{\varepsilon'''}(A|BC)_\rho + H_{\max}^{\varepsilon''}(B|C)_\rho + g\,,$$
$$H_{\max}^{\varepsilon'}(AB|C)_\rho \geq H_{\min}^{\varepsilon''}(A|BC)_\rho + H_{\min}^{\varepsilon'''}(B|C)_\rho - g\,,$$

where f does not grow more than of the order $\log 1/e$ when $e = \varepsilon - \varepsilon' - 2\varepsilon''$ is small, and g is smaller than 6 for $\varepsilon' + 2\varepsilon'' + \varepsilon''' < 1/5$. We note that, in typical applications, we would choose the smoothing parameters so that the correction terms f and g are small compared to the typical values of the smooth entropies.

The fact that generalized chain inequalities hold for smooth min- and max-entropy is not only important for establishing a complete entropy calculus, analogous to that for the von Neumann entropy. They are also crucial for applications. However, until now, only special cases of these inequalities have been known, except for the first pair, which has been derived in [11]. In the present paper we provide proofs for the remaining relations. In fact, since smooth min- and max-entropy obey a duality relation similar to that of von Neumann entropy, $H_{\min}^{\varepsilon}(A|B) = -H_{\max}^{\varepsilon}(A|C)$ (see Lemma 5), the paired inequalities above imply each other. It will therefore suffice to prove only one inequality of each pair.

The paper is organized as follows. In the next section we introduce the notation, terminology, and basic definitions. In particular, we define the (smooth) min- and max-entropy measures and outline some of their basic features. In Section III we derive alternative expressions for the max-entropy based on semidefinite programming duality. While these expressions may be of independent interest, they will be used in Section IV, which is devoted to the statement and proofs of the generalized chain rules.

II. Mathematical Preliminaries

A. Notation and basic definitions

Throughout this paper we focus on finite dimensional Hilbert spaces. Hilbert spaces corresponding to different physical systems are distinguished by different capital Latin letters as subscript $\mathcal{H}_A, \mathcal{H}_B$ etc. The tensor product of $\mathcal{H}_A$ and $\mathcal{H}_B$ is designated in short by $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$.

The set of linear operators from $\mathcal{H}_A$ to $\mathcal{H}_B$ is denoted by $\mathcal{L}(\mathcal{H}_A, \mathcal{H}_B)$. The space of linear operators acting on the Hilbert space $\mathcal{H}$ is denoted by $\mathcal{L}(\mathcal{H})$ and the subset of $\mathcal{L}(\mathcal{H})$ containing the Hermitian operators on $\mathcal{H}$ is denoted by $\mathrm{Herm}(\mathcal{H})$. Note that $\mathrm{Herm}(\mathcal{H})$ endowed with the Hilbert-Schmidt inner product $\langle X, Y\rangle := \mathrm{tr}(X^\dagger Y)$, $X, Y \in \mathrm{Herm}(\mathcal{H})$, is a Hilbert space. Given an operator $R \in \mathrm{Herm}(\mathcal{H})$, we write $R \geq 0$ if and only if R is positive semi-definite and $R > 0$ if and only if it is positive definite. Furthermore, let $\mathcal{S}_{\leq}(\mathcal{H})$ and $\mathcal{S}_{=}(\mathcal{H})$ denote the sets of sub-normalized and normalized positive semi-definite *density operators* with $\mathrm{tr}\,\rho \leq 1$ and $\mathrm{tr}\,\rho = 1$, respectively.

We generalize the notion of inequality to Hermitian operators in the following way: Let $R, S \in \mathrm{Herm}(\mathcal{H})$, then we write $R \geq S$, respectively $R > S$ if and only if $R - S$ is positive semi-definite, respectively positive definite.

Given an operator R, the operator norm of R is denoted by $\|R\|_\infty$ and is equal to the highest singular value of R. The trace norm of R is given by $\|R\|_1 := \mathrm{tr}[\sqrt{R^\dagger R}]$. The fidelity between two states ρ, $\sigma \in \mathcal{S}_{\leq}(\mathcal{H})$ is defined as $F(\rho, \sigma) := \|\sqrt{\rho}\sqrt{\sigma}\|_1$.

For multipartite operators on product spaces $\mathcal{H}_{AB}$ we will use subscripts to denote the space on which they act (e.g. S_{AB} for an operator on $\mathcal{H}_{AB}$). Given a multipartite operator $S_{AB} \in \mathcal{L}(\mathcal{H}_{AB})$, the corresponding reduced operator on $\mathcal{H}_A$ is defined by $S_A := \mathrm{tr}_B[S_{AB}]$ where tr_B denotes the partial trace operator on the subsystem $\mathcal{H}_B$. Given a multipartite operator S_{AB} and the corresponding marginal operator S_A, we call S_{AB} an *extension* of S_A. We omit identities from expressions which involve multipartite operators whenever mathematically meaningful expressions can be obtained by tensoring the corresponding identities to the operators.

B. Smooth Min- and Max-Entropies

In the following we successively give the definitions of the non-smooth min- and max-entropies and their smooth versions [3], [5].

Definition 1. *Let $\rho_{AB} \in \mathcal{S}_{\leq}(\mathcal{H}_{AB})$, then the min-entropy of A conditioned on B of ρ_{AB} is defined as*

$$H_{\min}(A|B)_\rho := \max_{\sigma_B \in \mathcal{S}_{\leq}(\mathcal{H}_B)} H_{\min}(A|B)_{\rho|\sigma}, \quad \textit{where}$$
$$H_{\min}(A|B)_{\rho|\sigma} := \sup\big\{\lambda \in \mathbb{R} : \rho_{AB} \leq 2^{-\lambda}\,\mathbb{I}_A \otimes \sigma_B\big\}. \quad (2)$$

Note that $H_{\min}(A|B)_{\rho|\sigma}$ is finite if and only if $\mathrm{supp}(\rho_B) \subseteq \mathrm{supp}(\sigma_B)$ and divergent otherwise.

Definition 2. *Let $\rho_{AB} \in \mathcal{S}_{\leq}(\mathcal{H}_{AB})$, then the max-entropy of A conditioned on B of ρ_{AB} is defined as*

$$H_{\max}(A|B)_\rho := \max_{\sigma_B \in \mathcal{S}_{\leq}(\mathcal{H}_B)} H_{\max}(A|B)_{\rho|\sigma}, \quad \textit{where}$$
$$H_{\max}(A|B)_{\rho|\sigma} := \log F(\rho_{AB}, \mathbb{I}_A \otimes \sigma_B)^2. \quad (3)$$

The maximum in (2) and (3) is achieved at $\mathcal{S}_{=}(\mathcal{H}_B)$. The ε-smooth min- and max-entropies of a state ρ can be understood as an optimization of the corresponding non-smooth quantities over a set of states ε-close to ρ. We use the *purified distance* to quantify the ε-closeness of states.

Definition 3. *Let ρ, $\sigma \in \mathcal{S}_{\leq}(\mathcal{H})$. Then the purified distance between ρ and σ is defined by*

$$P(\rho, \sigma) := \sqrt{1 - \bar{F}(\rho, \sigma)^2}, \quad \textit{where} \quad (4)$$
$$\bar{F}(\rho, \sigma) := F(\rho, \sigma) + \sqrt{\big(1 - \mathrm{tr}\,\rho\big)\big(1 - \mathrm{tr}\,\sigma\big)} \quad (5)$$

is the generalized fidelity.

From now on, when two states $\rho, \sigma \in \mathcal{S}_{\leq}(\mathcal{H})$ are said to be ε-close we mean $P(\rho,\sigma) \leq \varepsilon$ and denote this by $\rho \approx_\varepsilon \sigma$. Some of the basic properties of the purified distance are reviewed in Appendix B, but for a more comprehensive treatment we refer to [7]. With that convention we are ready to introduce a smoothed version of the min- and max-entropies [3].

Definition 4. *Let $\varepsilon \geq 0$, $\rho_{AB} \in \mathcal{S}_{\leq}(\mathcal{H}_{AB})$. Then the ε-smooth min-entropy of A conditioned on B of ρ_{AB} is defined as*

$$H_{\min}^{\varepsilon}(A|B)_\rho := \max_{\tilde{\rho}} H_{\min}(A|B)_{\tilde{\rho}} \tag{6}$$

and the ε-smooth max-entropy of A conditioned on B of ρ_{AB} is defined as

$$H_{\max}^{\varepsilon}(A|B)_\rho := \min_{\tilde{\rho}} H_{\max}(A|B)_{\tilde{\rho}} \tag{7}$$

where the maximum and the minimum range over all sub-normalized states $\tilde{\rho}_{AB} \approx_\varepsilon \rho_{AB}$.

The smooth min- and max-entropies are dual to each other in the following sense [7]:

Lemma 5. *Let $\varepsilon \geq 0$, $\rho_{AB} \in \mathcal{S}_{\leq}(\mathcal{H}_{AB})$ and $\rho_{ABC} \in \mathcal{S}_{\leq}(\mathcal{H}_{ABC})$ any purification of ρ_{AB}. Then,*

$$H_{\max}^{\varepsilon}(A|B)_\rho = -H_{\min}^{\varepsilon}(A|C)_\rho. \tag{8}$$

Finally, the smooth min-entropy is upper-bounded by the smooth max-entropy as shown by the following lemma whose proof is deferred to Appendix A:

Lemma 6. *Let $\rho_{AB} \in \mathcal{S}_{\leq}(\mathcal{H}_{AB})$ and $\varepsilon, \varepsilon' \geq 0$ such that $\varepsilon + \varepsilon' + 2\sqrt{1-\operatorname{tr}\rho_{AB}} < 1$. Then,*

$$\begin{aligned} H_{\min}^{\varepsilon'}(A|B)_\rho \leq\ & H_{\max}^{\varepsilon}(A|B)_\rho \\ & + \log\left(\frac{1}{1-(\varepsilon+\varepsilon'+2\sqrt{1-\operatorname{tr}\rho})^2}\right). \end{aligned} \tag{9}$$

C. Semidefinite Programming

This subsection is devoted to the duality theory of semi-definite programs (SDPs). We will present the subject as given in [13] and especially in [14] but will restrict the discussion to the special case which is of interest in this work.
A semidefinite program over the Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ is a triple $(\mathcal{F}, R_A, S_B)$, $\mathcal{F} \in \mathcal{L}(\mathrm{Herm}(\mathcal{H}_A), \mathrm{Herm}(\mathcal{H}_B))$, $R_A \in \mathrm{Herm}(\mathcal{H}_A)$ and $S_B \in \mathrm{Herm}(\mathcal{H}_B)$, which is associated with the following two optimization problems:

PRIMAL PROBLEM:		DUAL PROBLEM:	
minimize:	$\operatorname{tr}[R_A X_A]$	maximize:	$\operatorname{tr}[S_B Y_B]$
subject to:	$\mathcal{F}(X_A) \geq S_B$	subject to:	$\mathcal{F}^\dagger(Y_B) \leq R_A$
	$X_A \geq 0$		$Y_B \geq 0$

where $X_A \in \mathrm{Herm}(\mathcal{H}_A)$ and $Y_B \in \mathrm{Herm}(\mathcal{H}_B)$ are variables. $X_A \geq 0$ and $Y_B \geq 0$ such that $\mathcal{F}(X_A) \geq S_B$ and $\mathcal{F}^\dagger(Y_B) \leq R_A$, respectively, are called *primal feasible plan* and *dual feasible plan*, respectively. We also denote the solutions to the primal and dual problems by

$$\gamma := \inf\big\{\operatorname{tr}[R_A X_A] : X_A \text{ is a primal feasible plan}\big\},$$

$$\delta := \sup\big\{\operatorname{tr}[S_B Y_B] : Y_B \text{ is a dual feasible plan}\big\}.$$

The values $X_A \geq 0$ and $Y_B \geq 0$ satisfying $\operatorname{tr}[R_A X_A] = \gamma$ and $\operatorname{tr}[S_B Y_B] = \delta$ are called *primal optimal plan*, respectively *dual optimal plan*.
According to the *weak duality theorem* $\gamma \geq \delta$. The difference $\gamma - \delta$ is called *duality gap*. The following theorem called *Slater's condition* establishes an easy-to-check condition under which the duality gap vanishes, that is, $\gamma = \delta$.

Theorem 7. *Let γ and δ be defined as above and $(\mathcal{F}, R_A, S_A)$ with $R_A \in \mathrm{Herm}(\mathcal{H}_A)$ and $S_B \in \mathrm{Herm}(\mathcal{H}_B)$ a semi-definite program. Then the following two implications hold:*
(i)[Strict dual feasibility] Suppose γ is finite and that there exists an operator $Y_B > 0$ such that $\mathcal{F}^\dagger(Y_B) < R_A$. Then $\gamma = \delta$.
(ii) [Strict primal feasibility] Suppose that δ is finite and that there exists an operator $X_A > 0$ such that $\mathcal{F}(X_A) > S_B$. Then $\gamma = \delta$.

III. New Expressions and Bounds for the Smooth Max-Entropy

In the following, we give alternative expressions for $H_{\max}(A|B)_{\rho|\sigma}$ and $H_{\max}(A|B)_\rho$ based on the analysis of their corresponding SDPs. Then, we prove inequalities relating these entropies with a new entropic measure that turns out to be a useful tool for proving the chain rules.

A. New Expressions via SDP Duality

Lemma 8. *Let $\rho_{AB} \in \mathcal{S}_{\leq}(\mathcal{H}_{AB})$, $\sigma_B \in \mathcal{S}_{\leq}(\mathcal{H}_B)$ and let ρ_{ABC} be a purification of ρ_{AB} on an auxiliary Hilbert space $\mathcal{H}_C$. Then the max-entropy of A conditioned on B of ρ_{AB} relative to σ_B is given by*

$$H_{\max}(A|B)_{\rho|\sigma} = \log \min_{Z_{AB}} \operatorname{tr}[(\mathbb{I}_A \otimes \sigma_B) Z_{AB}], \tag{10}$$

where the minimum ranges over all $Z_{AB} \in \mathcal{P}(\mathcal{H}_{AB})$ with $\rho_{ABC} \leq Z_{AB} \otimes \mathbb{I}_C$.

Proof: Uhlmann's theorem [15] tells us that the fidelity can be expressed as a maximization of the overlap of purifications in which the optimization goes over one purification only. In particular, if ρ_{ABC} is a purification of ρ_{AB}, then

$$\begin{aligned} 2^{H_{\max}(A|B)_{\rho|\sigma}} &= F(\rho_{AB}, \mathbb{I}_A \otimes \sigma_B)^2 \\ &= \max_{\substack{X_{ABC} \geq 0 \\ \operatorname{tr}_C[X_{ABC}] = \mathbb{I}_A \otimes \sigma_B}} F(\rho_{ABC}, X_{ABC})^2 \\ &= \max_{\substack{X_{ABC} \geq 0 \\ \operatorname{tr}_C[X_{ABC}] = \mathbb{I}_A \otimes \sigma_B}} \operatorname{tr}[\rho_{ABC} X_{ABC}], \end{aligned} \tag{11}$$

where we optimize over all rank one extensions X_{ABC} of $\mathbb{I}_A \otimes \sigma_B$. Notice that instead of optimizing over pure states only we can let the maximization range over all positive semidefinite operators X_{ABC} since by Uhlmann's theorem we can always pick up large enough purifying system $\mathcal{H}_C$ such that there exists an optimal rank one X_{ABC}. Furthermore, for any positive semidefinite operator X_{ABC} with $\operatorname{tr}_C[X_{ABC}] \leq \mathbb{I}_A \otimes \sigma_B$ we can define an operator

$$\bar{X}_{ABC} := X_{ABC} + Y_C \otimes \big(\mathbb{I}_A \otimes \sigma_B - \operatorname{tr}_C X_{ABC}\big),$$

with Y_C an arbitrary element of $\mathcal{S}_=(\mathcal{H}_C)$. By construction it is constrained by $\mathrm{tr}_C\, \bar{X}_{ABC} = \mathbb{I}_A \otimes \sigma_B$ and also satisfies

$$\mathrm{tr}\big[\bar{X}_{ABC}\rho_{ABC}\big] \geq \mathrm{tr}\big[X_{ABC}\rho_{ABC}\big].$$

Hence, in (11) we can take the maximum over the set of all nonnegative operators X_{ABC} whose partial trace $\mathrm{tr}_C\, X_{ABC}$ is bounded by $\mathbb{I}_A \otimes \sigma_B$ (in spite of being equal to $\mathbb{I}_A \otimes \sigma_B$). The SDP for the conditional max-entropy of a state $\rho_{AB} \in \mathcal{S}_\leq(\mathcal{H}_{AB})$ relative to a given $\sigma_B \in \mathcal{S}_\leq(\mathcal{H}_{AB})$ is as follows:

	PRIMAL PROBLEM:		DUAL PROBLEM:
minimum:	$\mathrm{tr}[(\mathbb{I}_A \otimes \sigma_B) Z_{AB}]$	maximum:	$\mathrm{tr}[X_{ABC}\rho_{ABC}]$
subject to:	$Z_{AB} \otimes \mathbb{I}_C \geq \rho_{ABC}$	subject to:	$\mathrm{tr}_C[X_{ABC}] \leq \mathbb{I}_A \otimes \sigma_B$
	$Z_{AB} \geq 0.$		$X_{ABC} \geq 0$

where Z_{AB} is a primal variable and X_{ABC} a dual variable, respectively. Since the space in the dual problem over which one is optimizing, is closed and bounded, it is compact by the Weierstrass theorem. Hence, the dual optimal plan is finite. Furthermore, the operator $\bar{Z}_{AB} = 2\|\rho_{ABC}\|_\infty \mathbb{I}_{AB} > 0$ satisfies Slater's strict primal feasibility condition $2\|\rho_{ABC}\|_\infty \mathbb{I}_{ABC} - \rho_{ABC} > 0$ and thus the duality gap between the primal and dual optimization problems vanishes. ∎

Next, we write out the SDP for $H_{\max}(A|B)_\rho$ and explore the duality gap between the optimization problems.

Lemma 9. *Let $\rho_{AB} \in \mathcal{S}_\leq(\mathcal{H}_{AB})$ and let ρ_{ABC} be a purification of ρ_{AB} on an auxiliary Hilbert space $\mathcal{H}_C$. Then the max-entropy of A conditioned on B of ρ_{AB} is given by*

$$H_{\max}(A|B)_\rho := \log \min_{Z_{AB}} \|Z_B\|_\infty, \tag{12}$$

where the minimum ranges over all $Z_{AB} \in \mathcal{P}(\mathcal{H}_{AB})$ with $\rho_{ABC} \leq Z_{AB} \otimes \mathbb{I}_C$.

Proof: The only thing that changes with respect to the SDP in Lemma 8 is that σ_B is no longer fixed but it becomes a dual variable. Thus the SDP for $H_{\max}(A|B)_\rho$ reads:

	PRIMAL PROBLEM:		DUAL PROBLEM:
minimum:	λ	maximum:	$\mathrm{tr}[X_{ABC}\rho_{ABC}]$
subject to:	$Z_{AB} \otimes \mathbb{I}_C \geq \rho_{ABC}$	subject to:	$\mathrm{tr}_C[X_{ABC}] \leq \mathbb{I}_A \otimes \sigma_B$
	$\lambda \mathbb{I}_B \geq \mathrm{tr}_A[Z_{AB}]$		$\mathrm{tr}[\sigma_B] \leq 1$
	$Z_{AB} \geq 0,\ \lambda \geq 0$		$X_{ABC} \geq 0,\ \sigma_B \geq 0$

where λ and Z_{AB} are primal variables and σ_B and X_{ABC} dual variables. Obviously, the optimal λ is equal to the largest eigenvalue of Z_B. Hence, the above program may be rewritten in the form:

	PRIMAL PROBLEM:		DUAL PROBLEM:
minimum:	$\|Z_B\|_\infty$	maximum:	$\mathrm{tr}[X_{ABC}\rho_{ABC}]$
subject to:	$Z_{AB} \otimes \mathbb{I}_C \geq \rho_{ABC}$	subject to:	$\mathrm{tr}_C[X_{ABC}] \leq \mathbb{I}_A \otimes \sigma_B$
	$Z_{AB} \geq 0$		$\mathrm{tr}[\sigma_B] \leq 1$
			$X_{ABC} \geq 0,\ \sigma_B \geq 0$

In the dual problem we are optimizing over compact sets, thus there exists a finite dual optimal plan. Furthermore, $\bar{Z}_{AB} = 2\|\rho_{ABC}\|_\infty \mathbb{I}_{AB} > 0$ and $\bar{\lambda} = 2\|\bar{Z}_B\|_\infty > 0$ satisfy Slater's strict primal feasibility condition $\bar{Z}_{AB} \otimes \mathbb{I}_C > \rho_{ABC}$ and $\bar{\lambda}\mathbb{I}_B > \mathrm{tr}_A[\bar{Z}_{AB}]$ which implies a zero duality gap. ∎

Note that one can always write the operator norm of Z_B as

$$\|Z_B\|_\infty = \max_{\sigma_B} \mathrm{tr}[\sigma_B Z_B] = \max_{\sigma_B} \mathrm{tr}[(\mathbb{I}_A \otimes \sigma_B) Z_{AB}],$$

where the maximum ranges over all $\sigma_B \in \mathcal{S}_\leq(\mathcal{H}_B)$. Expression (12) then acquires the form

$$H_{\max}(A|B)_\rho = \log \min_{\rho_{ABC} \leq Z_{AB} \otimes \mathbb{I}_C} \max_{\sigma_B} \mathrm{tr}[(\mathbb{I}_A \otimes \sigma_B) Z_{AB}]. \tag{13}$$

On the other hand from the vanishing of the duality gap in the SDP of $H_{\max}(A|B)_{\rho|\sigma}$ it follows that

$$\log F(\rho_{AB}, \mathbb{I}_A \otimes \sigma_B)^2 = \log \min_{Z_{AB}} \mathrm{tr}[(\mathbb{I}_A \otimes \sigma_B) Z_{AB}]$$

which after maximization of the left- and the right-hand sides over $\sigma_B \in \mathcal{S}_\leq(\mathcal{H}_B)$ implies

$$H_{\max}(A|B)_\rho = \log \max_{\sigma_B} \min_{Z_{AB}} \mathrm{tr}[(\mathbb{I}_A \otimes \sigma_B) Z_{AB}].$$

Therefore, the operations $\min$ and $\max$ in (13) commute. In the following we restate Lemma 6 and prove it using the derived SDPs for the non-smooth max-entropy.

Henceforth, we will use (3), (10) and (12) and (13) as interchangeable expressions for the conditional max-entropy and the conditional relative max-entropy, respectively.

B. A Bound on the Relative Conditional Entropy

We prove an important technical lemma, which we later use to derive one of the chain rules.

Lemma 10. *Let $\rho_{AB} \in \mathcal{S}_\leq(\mathcal{H}_{AB})$, $\rho'_{AB} \approx_{\varepsilon'} \rho_{AB}$ and $\varepsilon > 0$. Then there exists a state $\tilde{\rho}_{AB} \approx_{\varepsilon+\varepsilon'} \rho'_{AB}$ such that*

$$H_{\max}(A|B)_{\tilde{\rho}} \leq H_{\max}(A|B)_{\rho|\rho'} + \log\left(\frac{1}{1-\sqrt{1-\varepsilon^2}}\right). \tag{14}$$

Proof: Let $\tilde{Z}_{AB}$ be an optimal primal plan for the semidefinite program for $H_{\max}(A|B)_{\rho|\rho'}$ and Π_B be the minimum rank projector onto the smallest eigenvalues of the reduced operator $\tilde{Z}_B$ such that $\mathrm{tr}[\Pi_B^\perp \rho'_B] \leq 1 - \sqrt{1-\varepsilon^2}$ where $\Pi_B^\perp$ is the orthogonal complement of Π_B and let $\tilde{\rho}_{AB} := \Pi_B \rho_{AB} \Pi_B$. By Equation (12), we can write

$$\begin{aligned} 2^{H_{\max}(A|B)_{\tilde{\rho}}} &= \min_{\tilde{\rho}_{ABC} \leq \tilde{Z}_{AB} \otimes \mathbb{I}_C} \|Z_B\|_\infty \\ &\leq \|\Pi_B \tilde{Z}_B \Pi_B\|_\infty, \end{aligned}$$

where we used the fact that $\rho_{ABC} \leq \tilde{Z}_{AB} \otimes \mathbb{I}_C$ implies $\tilde{\rho}_{ABC} \leq \Pi_B \tilde{Z}_{AB} \Pi_B \otimes \mathbb{I}_C$. Let Π'_B be the projector onto the largest eigenvalue of $\Pi_B \tilde{Z}_B \Pi_B$. Since Π'_B and $\Pi_B^\perp$ project on eigenvectors of $\tilde{Z}_B$, they commute with $\tilde{Z}_B$. Then,

$$\begin{aligned} \|\Pi_B \tilde{Z}_B \Pi_B\|_\infty &= \mathrm{tr}[\Pi'_B \tilde{Z}_B] \\ &= \min_{\mu_B} \frac{\mathrm{tr}[\mu_B \tilde{Z}_B]}{\mathrm{tr}[\mu_B]}, \end{aligned} \tag{15}$$

where the minimization is over all positive operators in the support of $\Pi_B^\perp + \Pi'_B$. Fixing $\mu_B = (\Pi_B^\perp + \Pi'_B)\rho'_B(\Pi_B^\perp + \Pi'_B)$,

we obtain the following upper bound for (15):

$$\begin{aligned}\|\Pi_B\tilde{Z}_B\Pi_B\|_\infty &\leq \frac{\mathrm{tr}[(\Pi_B^\perp+\Pi_B')\rho_B'(\Pi_B^\perp+\Pi_B')\tilde{Z}_B]}{\mathrm{tr}[(\Pi_B^\perp+\Pi_B')\rho_B']}\\ &= \frac{\mathrm{tr}[(\Pi_B^\perp+\Pi_B')\tilde{Z}_B^{1/2}\rho_B'\tilde{Z}_B^{1/2}]}{\mathrm{tr}[(\Pi_B^\perp+\Pi_B')\rho_B']}\\ &\leq \frac{\mathrm{tr}[\rho_B'\tilde{Z}_B]}{\mathrm{tr}[(\Pi_B^\perp+\Pi_B')\rho_B']}\\ &\leq 2^{H_{\max}(A|B)_{\rho|\rho'}}\frac{1}{1-\sqrt{1-\varepsilon^2}},\end{aligned}$$

where we used Equation (10) and the fact that $\mathrm{tr}[(\Pi_B^\perp+\Pi_B')\rho_B'] \geq 1-\sqrt{1-\varepsilon^2}$ by definition of $\Pi_B^\perp$. Then, taking the logarithm on both sides yields (14).

Finally, the proof is concluded by the upper bound

$$\begin{aligned}&P(\tilde{\rho}_{AB},\rho_{AB}') = P(\Pi_B\rho_{AB}\Pi_B,\rho_{AB}')\\ &\leq P(\Pi_B\rho_{AB}\Pi_B,\Pi_B\rho_{AB}'\Pi_B)+P(\Pi_B\rho_{AB}'\Pi_B,\rho_{AB}')\\ &\leq P(\rho_{AB},\rho_{AB}')+\sqrt{2\,\mathrm{tr}[\Pi_B^\perp\rho_{AB}']-(\mathrm{tr}[\Pi_B^\perp\rho_{AB}'])^2}\\ &\leq \varepsilon'+\varepsilon\end{aligned}$$

where we use Inequality (34) and the fact that the function $\sqrt{2t-t^2}$ is monotonously increasing in the interval $[0,1]$. ■

C. The ε-Smooth S-Entropy

To prove the chain rules, we also need an auxiliary entropy measure called ε-smooth S-entropy[2] whose definition and basic properties are given in the following.

We assume that $\rho_{AB} \in \mathcal{S}_\leq(\mathcal{H}_{AB})$ and $\sigma_B \in \mathcal{S}_\leq(\mathcal{H}_B)$ with $\mathrm{supp}(\rho_B) \subseteq \mathrm{supp}(\sigma_B)$ and denote for every $\lambda \in \mathbb{R}$ the projector onto the eigenspace corresponding to the negative eigenvalues of the operator $2^\lambda\rho_{AB}-\sigma_B$ by P_{AB}^λ.

Definition 11. *Let $\varepsilon > 0$. Then the ε-smooth S-entropy of A conditioned on B of ρ_{AB} relative to σ_B is defined as*

$$S^\varepsilon(A|B)_{\rho|\sigma} := \inf\{\lambda\in\mathbb{R} : \mathrm{tr}[P_{AB}^\lambda\rho_{AB}] \leq \varepsilon\}. \qquad (16)$$

We would like to upper bound this new quantity in terms of the max-entropy. In order to achieve this we prove the following technical lemma:

Lemma 12. *Let $\varepsilon > 0$ and $\lambda_{\inf} \in \mathbb{R}$ the infimum as in Definition 11, then there exists a number $\lambda \in \mathbb{R}$ such that $\lambda \geq \lambda_{\inf}$ and $\mathrm{tr}[P_{AB}^\lambda\rho_{AB}] \geq \varepsilon$.*

Proof: From the assumption $\mathrm{supp}(\rho_{AB}) \subseteq \mathrm{supp}(\mathbb{I}_A\otimes\sigma_B)$ it follows that one can always find a sufficiently small real number λ so that the operator $2^\lambda\rho_{AB}-\sigma_B$ becomes negative definite on $\mathcal{H}_A\otimes\mathrm{supp}(\sigma_B)$. For any such λ we trivially have $\mathrm{tr}[P_{AB}^\lambda\rho_{AB}] \geq \varepsilon$. Define $\lambda_{\sup} := \sup\{\lambda\in\mathbb{R} : \mathrm{tr}[P_{AB}^\lambda\rho_{AB}] \geq \varepsilon\}$ and assume that $\lambda_{\sup} < \lambda_{\inf}$. Then for every $\lambda \in (\lambda_{\sup},\lambda_{\inf})$ we would have $\mathrm{tr}[P_{AB}^\lambda\rho_{AB}] < \varepsilon$ which however contradicts the assumption that $\lambda_{\inf}$ is an infimum. Therefore, we conclude that $\lambda_{\inf} \leq \lambda_{\sup}$. Thus one can always find a sufficiently small $\delta \geqslant 0$ such that the number $\lambda := \lambda_{\sup}-\delta \geq \lambda_{\inf}$ and $\mathrm{tr}[P_{AB}^\lambda\rho_{AB}] \geq \varepsilon$ holds which concludes the proof. ■

The next lemma gives the upper bound of the ε-smooth S-entropy in terms of the max-entropy.

Lemma 13. *Let $\rho_{AB} \in \mathcal{S}_\leq(\mathcal{H}_{AB})$, $\sigma_B \in \mathcal{S}_\leq(\mathcal{H}_B)$ and $\varepsilon > 0$. Then,*

$$S^\varepsilon(A|B)_{\rho|\sigma} \leq H_{\max}(A|B)_{\rho|\sigma} + \log\left(\frac{1}{\varepsilon^2}\right). \qquad (17)$$

Proof: Let $\lambda_{\inf} \in \mathbb{R}$ be the infimum in Definition 11, that is, $\lambda_{\inf} = S^\varepsilon(A|B)_{\rho|\sigma}$, λ be as in Lemma 12 and $P_{AB}^\pm$ denote the projector onto the positive/negative eigenvalues of $\rho_{AB}-2^{-\lambda}\sigma_B$, respectively. Then, a straightforward computation yields

$$\begin{aligned}2^{\frac{1}{2}H_{\max}(A|B)_{\rho|\sigma}-\frac{1}{2}S^\varepsilon(A|B)_{\rho|\sigma}} &= \|\sqrt{\rho_{AB}}\sqrt{\sigma_B}\|_1 2^{-\frac{1}{2}\lambda_{\inf}}\\ &\geq \mathrm{tr}[\sqrt{\rho_{AB}}\sqrt{2^{-\lambda_{\inf}}\sigma_B}]\\ &\geq \mathrm{tr}[\sqrt{\rho_{AB}}\sqrt{2^{-\lambda}\sigma_B}]\\ &\geq \mathrm{tr}[P_{AB}^+2^{-\lambda}\sigma_B + P_{AB}^-\rho_{AB}]\\ &\geq \mathrm{tr}[P_{AB}^-\rho_{AB}]\\ &\geq \varepsilon, \qquad (18)\end{aligned}$$

where we applied Corollary 18 and used the fact that P_{AB}^- is identical with the projector P_{AB}^λ. Taking the logarithm on both sides of (18) and rearranging the terms we obtain (17). ■

IV. Main Results

This section contains the main result of this paper: a derivation of the previously unknown chain rules for smooth min- and max-entropies. To simplify presentation hereafter, we introduce the function

$$f:\ \varepsilon \mapsto \log\frac{1}{1-\sqrt{1-\varepsilon^2}}$$

that appears as an error term in the chain rules. It vanishes as $\varepsilon \to 1$ and grows logarithmically in $\frac{1}{\varepsilon}$ when $\varepsilon \to 0$.

As remarked in the introduction, the explicit form of two of the chain rules has already been derived in [11], namely

$$H_{\min}^{\varepsilon+\varepsilon'+2\varepsilon''}(AB|C)_\rho \geq H_{\min}^{\varepsilon''}(A|BC)_\rho + H_{\min}^{\varepsilon'}(B|C)_\rho - f(\varepsilon)$$

and its dual. Here we provide proofs for the remaining three pairs of chain rules. Due to the smooth duality relation (8) it is enough to prove only one of each pair.

Theorem 14. *Let $\varepsilon > 0$, $\varepsilon', \varepsilon'' \geq 0$ and $\rho_{ABC} \in \mathcal{S}_\leq(\mathcal{H}_{ABC})$. Then,*

$$H_{\min}^{\varepsilon'}(AB|C)_\rho \leq H_{\min}^{\varepsilon+\varepsilon'+2\varepsilon''}(A|BC)_\rho + H_{\max}^{\varepsilon''}(B|C)_\rho + 2f(\varepsilon). \qquad (19)$$

Proof: Let $\rho_{ABC}' \approx_{\varepsilon'} \rho_{ABC}$, $\rho_{BC}'' \approx_{\varepsilon''} \rho_{BC}$ and $\sigma_C \in \mathcal{S}_\leq(\mathcal{H}_C)$ such that

$$\begin{aligned}H_{\min}(AB|C)_{\rho'} &= H_{\min}^{\varepsilon'}(AB|C)_\rho,\\ H_{\max}(B|C)_{\rho''} &= H_{\max}^{\varepsilon''}(B|C)_\rho,\end{aligned}$$

[2]The idea for this new entropy measure was originally proposed by Robert König.

and

$$\rho'_{ABC} \leq 2^{-H_{\min}(AB|C)_{\rho'}}\sigma_C = 2^{-H^{\varepsilon'}_{\min}(AB|C)_{\rho}}\sigma_C. \quad (20)$$

For every $\delta > 0$ there is a $\delta' \in (0, \delta]$ such that the projector P^{λ}_{BC} onto the negative eigenvalues of the operator $2^{\lambda}\rho''_{BC} - \sigma_C$ with $\lambda := S^{\tilde{\varepsilon}}(B|C)_{\rho''|\sigma} + \delta'$, $\tilde{\varepsilon} > 0$, satisfies the constraint $\mathrm{tr}[P^{\lambda}_{BC}\rho''_{BC}] \leq \tilde{\varepsilon}$ in Definition 11. If $P^{\lambda\perp}_{BC}$ is the orthogonal complement of P^{λ}_{BC}, we have

$$P^{\lambda\perp}_{BC}\sigma_C P^{\lambda\perp}_{BC} \leq 2^{\lambda} P^{\lambda\perp}_{BC}\rho''_{BC}P^{\lambda\perp}_{BC}. \quad (21)$$

A conjugation of (20) with $P^{\lambda\perp}_{BC}$ together with (21) yields

$$P^{\lambda\perp}_{BC}\rho'_{ABC}P^{\lambda\perp}_{BC} \leq 2^{-H^{\varepsilon'}_{\min}(AB|C)_{\rho}+\lambda} P^{\lambda\perp}_{BC}\rho''_{BC}P^{\lambda\perp}_{BC},$$

which is equivalent to

$$H_{\min}(A|BC)_{P^{\lambda\perp}\rho' P^{\lambda\perp}|P^{\lambda\perp}\rho'' P^{\lambda\perp}} \geq H^{\varepsilon'}_{\min}(AB|C)_{\rho} - \lambda.$$

A subsequent optimization of the left-hand side over all $\mathcal{S}_{\leq}(\mathcal{H}_{BC})$ yields

$$H_{\min}(A|BC)_{P^{\lambda\perp}\rho' P^{\lambda\perp}} \geq H^{\varepsilon'}_{\min}(AB|C)_{\rho} - \lambda \quad (22)$$

Since ρ_{ABC} is an extension of ρ_{BC}, by Corollary 22 there exists an extension ρ''_{ABC} of ρ''_{BC} such that $P(\rho''_{ABC}, \rho_{ABC}) = P(\rho''_{BC}, \rho_{BC})$. Then Inequality (33) and Inequality (34) give us the following upper bound for the purified distance between $P^{\lambda\perp}_{BC}\rho'_{ABC}P^{\lambda\perp}_{BC}$ and ρ_{ABC}:

$$\begin{aligned} &P(P^{\lambda\perp}_{BC}\rho'_{ABC}P^{\lambda\perp}_{BC}, \rho_{ABC}) \\ &\qquad \leqslant P(P^{\lambda\perp}_{BC}\rho'_{ABC}P^{\lambda\perp}_{BC}, P^{\lambda\perp}_{BC}\rho_{ABC}P^{\lambda\perp}_{BC}) \\ &\qquad + P(P^{\lambda\perp}_{BC}\rho_{ABC}P^{\lambda\perp}_{BC}, P^{\lambda\perp}_{BC}\rho''_{ABC}P^{\lambda\perp}_{BC}) \\ &\qquad + P(P^{\lambda\perp}_{BC}\rho''_{ABC}P^{\lambda\perp}_{BC}, \rho'_{ABC}) \\ &\qquad + P(\rho''_{ABC}, \rho_{ABC}) \\ &\qquad \leq \sqrt{2\tilde{\varepsilon} - \tilde{\varepsilon}^2} + \varepsilon' + 2\varepsilon''. \end{aligned}$$

After smoothing the left-hand side of (22) and upper-bounding the term $S^{\tilde{\varepsilon}}(B|C)_{\rho''|\sigma}$ on the right-hand side of (22) by $H_{\max}(B|C)_{\rho''|\sigma}$ in accordance with Lemma 13 and subsequently optimizing it over $\mathcal{S}_{\leq}(\mathcal{H}_C)$, we obtain

$$\begin{aligned} H^{\varepsilon'}_{\min}(AB|C)_{\rho} \leq H^{\sqrt{2\tilde{\varepsilon}-\tilde{\varepsilon}^2}+\varepsilon'+2\varepsilon''}_{\min}(A|BC)_{\rho} + H^{\varepsilon''}_{\max}(B|C)_{\rho} \\ + \log\frac{1}{\tilde{\varepsilon}^2} + \delta'. \end{aligned}$$

Finally, the substitution $\tilde{\varepsilon} := 1 - \sqrt{1-\varepsilon^2}$ leads to the chain rule (19) in the limit $\delta \longrightarrow 0$. ∎

Theorem 15. *Let $\varepsilon > 0$, $\varepsilon', \varepsilon'' \geq 0$ and $\rho_{ABC} \in \mathcal{S}_{\leq}(\mathcal{H}_{ABC})$. Then,*

$$H^{\varepsilon'}_{\min}(AB|C)_{\rho} \leq H^{\varepsilon''}_{\max}(A|BC)_{\rho} + H^{2\varepsilon+\varepsilon'+2\varepsilon''}_{\min}(B|C)_{\rho} + 3f(\varepsilon). \quad (23)$$

Proof: Let ρ_{ABCD} be a purification of ρ_{ABC}. If

$$H^{\varepsilon'}_{\max}(AB|D)_{\rho} \geq H^{2\varepsilon+\varepsilon'+2\varepsilon''}_{\max}(B|AD)_{\rho} + H^{\varepsilon''}_{\min}(A|D)_{\rho} - 3f(\varepsilon)$$

holds, then the chain rule follows by the duality relation (8). Let $\rho'_{ABD} \approx_{\varepsilon'} \rho_{ABD}$, $\rho''_{AD} \approx_{\varepsilon''} \rho_{AD}$ and $\sigma_D \in \mathcal{S}_{\leq}(\mathcal{H}_D)$ s.t.

$$\begin{aligned} H_{\max}(AB|D)_{\rho'} &= H^{\varepsilon'}_{\max}(AB|D)_{\rho}, \\ H_{\min}(A|D)_{\rho''} &= H^{\varepsilon''}_{\min}(A|D)_{\rho}, \end{aligned}$$

and

$$\rho''_{AD} \leq 2^{-H_{\min}(A|D)_{\rho''}}\sigma_D = 2^{-H^{\varepsilon''}_{\min}(A|D)_{\rho}}\sigma_D. \quad (24)$$

Again we use the fact that for every $\delta > 0$ there exists a $\delta' \in (0, \delta]$ such that for $\lambda := S^{\tilde{\varepsilon}}(AB|D)_{\rho'|\sigma} + \delta'$, $\tilde{\varepsilon} > 0$, the projector P^{λ}_{ABD} onto the negative eigenvalues of the operator $2^{\lambda}\rho'_{ABD} - \sigma_D$ satisfies the constraint $\mathrm{tr}[P^{\lambda}_{ABD}\rho'_{ABD}] \leq \tilde{\varepsilon}$ in Definition 11. If $P^{\lambda\perp}_{ABD}$ denotes the orthogonal complement of P^{λ}_{ABD}, then

$$2^{\lambda} P^{\lambda\perp}_{ABD}\rho'_{ABD}P^{\lambda\perp}_{ABD} \geq P^{\lambda\perp}_{ABD}\sigma_D P^{\lambda\perp}_{ABD}. \quad (25)$$

A conjugation of (24) with $P^{\lambda\perp}_{ABD}$ and a subsequent combination with (25) yields

$$2^{\lambda - H^{\varepsilon''}_{\min}(A|D)_{\rho}} P^{\lambda\perp}_{ABD}\rho'_{ABD}P^{\lambda\perp}_{ABD} \geq P^{\lambda\perp}_{ABD}\rho''_{AD}P^{\lambda\perp}_{ABD}. \quad (26)$$

Consider now the max-entropy

$$2^{H_{\max}(B|AD)_{P^{\lambda\perp}\rho' P^{\lambda\perp}|\rho''}} = \min_{\substack{Z_{ABD}\geq 0 \\ P^{\lambda\perp}_{ABD}\rho'_{ABCD}P^{\lambda\perp}_{ABD} \leq Z_{ABD}\otimes\mathbb{I}_C}} \mathrm{tr}[(\mathbb{I}_B \otimes \rho''_{AD})Z_{ABD}] \quad (27)$$

where ρ'_{ABCD} is a purification of ρ'_{ABD}. Making use of (26) and the inequality

$$P^{\lambda\perp}_{ABD}\rho'_{ABCD}P^{\lambda\perp}_{ABD} \leq P^{\lambda\perp}_{ABD} \otimes \mathbb{I}_C$$

and omitting the identity operator, we can upper-bound the right-hand side of (27) in the following way:

$$\begin{aligned} &\leq \mathrm{tr}[\rho''_{AD}P^{\lambda\perp}_{ABD}] \\ &\leq 2^{\lambda - H^{\varepsilon''}_{\min}(A|D)_{\rho}}\,\mathrm{tr}[P^{\lambda\perp}_{ABD}\rho'_{ABD}P^{\lambda\perp}_{ABD}] \\ &\leq 2^{\lambda - H^{\varepsilon''}_{\min}(A|D)_{\rho}}, \end{aligned}$$

where we use that the term $\mathrm{tr}[P^{\lambda\perp}_{ABD}\rho'_{ABD}P^{\lambda\perp}_{ABD}]$ is upper bounded by one. Taking the logarithm and substituting λ yields

$$H_{\max}(B|AD)_{P^{\lambda\perp}\rho' P^{\lambda\perp}|\rho''} \leq S^{\tilde{\varepsilon}}(AB|D)_{\rho'|\sigma} + \delta' - H^{\varepsilon''}_{\min}(A|D)_{\rho}.$$

A subsequent application of Lemma 13 implies

$$\begin{aligned} H_{\max}(B|AD)_{P^{\lambda\perp}\rho' P^{\lambda\perp}|\rho''} \leq H_{\max}(AB|D)_{\rho'} - H^{\varepsilon''}_{\min}(A|D)_{\rho} \\ + \delta' + \log\frac{1}{\tilde{\varepsilon}^2}, \end{aligned} \quad (28)$$

where the max-entropy term on the right-hand side has been optimized on $\mathcal{S}_{\leq}(\mathcal{H}_D)$. Consider now the left-hand side of (28). Corollary 22 guarantees the existence of an extension ρ''_{ABD} such that $P(\rho''_{AD}, \rho_{AD}) = P(\rho''_{ABD}, \rho_{ABD})$. Then, it follows that

$$\begin{aligned} P(P^{\lambda\perp}_{ABD}\rho'_{ABD}P^{\lambda\perp}_{ABD}, \rho''_{ABD}) &\leq P(P^{\lambda\perp}_{ABD}\rho'_{ABD}P^{\lambda\perp}_{ABD}, \rho'_{ABD}) \\ &\quad + P(\rho'_{ABD}, \rho''_{ABD}) \\ &\leq \sqrt{2\tilde{\varepsilon} - \tilde{\varepsilon}^2} + \varepsilon' + \varepsilon''. \end{aligned}$$

Thus, according to Lemma 10, there exists a state $\tilde{\rho}_{ABD} \approx_{\varepsilon+\sqrt{2\tilde{\varepsilon}-\tilde{\varepsilon}^2}+\varepsilon'+2\varepsilon''} \rho_{ABD}$ such that

$$\begin{aligned} H_{\max}(B|AD)_{\tilde{\rho}} \leq H^{\varepsilon'}_{\max}(AB|D)_{\rho} - H^{\varepsilon''}_{\min}(A|D)_{\rho} \\ + \delta' + \log\frac{1}{\tilde{\varepsilon}^2} + f(\varepsilon). \end{aligned}$$

Smoothing of the left-hand side and regrouping the terms in the last inequality yields

$$H_{\max}^{\varepsilon'}(AB|D)_\rho \geq H_{\max}^{\varepsilon+\sqrt{2\tilde{\varepsilon}-\tilde{\varepsilon}^2}+\varepsilon'+2\varepsilon''}(B|AD)_\rho + H_{\min}^{\varepsilon''}(A|D)_\rho - \delta' - \log\frac{1}{\tilde{\varepsilon}^2} - f(\varepsilon).$$

Finally, setting $\tilde{\varepsilon} := 1-\sqrt{1-\varepsilon^2}$, taking the limit $\delta \to 0$, and applying the duality relation for smooth entropies (8), we obtain chain rule (23). ∎

The last chain rule follows from chain rule (19) together with Lemma 6.

Corollary 16. *Let ε', ε'' $\varepsilon''' \geq 0$ and $\rho_{ABC} \in \mathcal{S}_\leq(\mathcal{H}_{ABC})$ such that $\varepsilon' + 2\varepsilon'' + \varepsilon''' < 1 - 2\sqrt{1-\operatorname{tr}\rho}$. Then,*

$$H_{\min}^{\varepsilon'}(AB|C)_\rho \leq H_{\max}^{\varepsilon'''}(A|BC)_\rho + H_{\max}^{\varepsilon''}(B|C)_\rho + g(\varepsilon',\varepsilon'',\varepsilon''',\operatorname{tr}\rho)\,, \tag{29}$$

where $g(\varepsilon',\varepsilon'',\varepsilon''',\operatorname{tr}\rho) :=$

$$\inf_\varepsilon \left\{2f(\varepsilon) + \log\left(\frac{1}{1-(\varepsilon+\varepsilon'+2\varepsilon''+\varepsilon'''+2\sqrt{1-\operatorname{tr}\rho})^2}\right)\right\},$$

and the infimum is taken in the range $0 < \varepsilon < 1-\varepsilon'-2\varepsilon''-\varepsilon'''-2\sqrt{1-\operatorname{tr}\rho}$.

Proof: Let $\varepsilon > 0$ be any smoothing parameter such that $\varepsilon < 1-\varepsilon'-2\varepsilon''-\varepsilon'''-2\sqrt{1-\operatorname{tr}\rho}$. Then, by Lemma 6, the smooth min-entropy term on the right-hand side of (19) is upper bounded by

$$H_{\max}^{\varepsilon'''}(A|BC)_\rho + \log\left(\frac{1}{1-(\varepsilon+\varepsilon'+2\varepsilon''+\varepsilon'''+2\sqrt{1-\operatorname{tr}\rho})^2}\right)$$

which immediately gives (29). ∎

In contrast to the previous chain rules, the last one leads to non-trivial results even if we apply it to non-smooth entropies. In particular, for a normalized state ρ_{ABC}, we find

$$H_{\min}(AB|C)_\rho \leq H_{\max}(A|BC)_\rho + H_{\max}(B|C)_\rho + 4\,.$$

Acknowledgments

This work was supported by the Swiss National Science Foundation (SNF) through the National Centre of Competence in Research Quantum Science and Technology and project No. 200020-135048, and by the European Research Council (ERC) via grant No. 258932. MT acknowledges support from the National Research Foundation (Singapore), and the Ministry of Education (Singapore).

Appendix A
Proof Lemma 6

Restatement of Lemma 6. *Let $\rho_{AB} \in \mathcal{S}_\leq(\mathcal{H}_{AB})$ and $\varepsilon, \varepsilon' \geq 0$ such that $\varepsilon + \varepsilon' + 2\sqrt{1-\operatorname{tr}\rho_{AB}} < 1$. Then,*

$$H_{\min}^{\varepsilon'}(A|B)_\rho \leq H_{\max}^{\varepsilon}(A|B)_\rho + \log\left(\frac{1}{1-(\varepsilon+\varepsilon'+2\sqrt{1-\operatorname{tr}\rho})^2}\right).$$

Proof: Define $\hat{\rho}_{AB} = \rho_{AB}/\operatorname{tr}(\rho_{AB})$. According to Lemma 5.2 in [16] there are embeddings $U : \mathcal{H}_A \longrightarrow \mathcal{H}_{A'}$ and $V : \mathcal{H}_B \longrightarrow \mathcal{H}_{B'}$ such that there exists a normalized state $\bar{\rho}_{A'B'} \approx_\varepsilon \hat{\rho}_{A'B'}$, where $\hat{\rho}_{A'B'} = (U\otimes V)\,\hat{\rho}_{AB}\left(U^\dagger \otimes V^\dagger\right)$, which minimizes the smooth max-entropy $H_{\max}^{\tilde{\varepsilon}}(A'|B')_{\hat{\rho}} = H_{\max}^{\tilde{\varepsilon}}(A|B)_{\hat{\rho}}$.

Consider now the quantity $2^{-H_{\min}^{\tilde{\varepsilon}+\tilde{\varepsilon}'}(A'|B')_{\bar{\rho}}}$. We are simultaneously minimizing over all $\sigma_{B'} \in \mathcal{S}_\leq(\mathcal{H}_{B'})$ and all states $\tilde{\rho}_{A'B'}$, that are $\tilde{\varepsilon}+\tilde{\varepsilon}'$-close to the normalized state $\bar{\rho}_{A'B'}$. By Uhlmann's theorem the latter constraint translates into $\operatorname{tr}[\tilde{\rho}_{A'B'C}\bar{\rho}_{A'B'C}] \geq 1-(\tilde{\varepsilon}+\tilde{\varepsilon}')^2$ where $\mathcal{H}_C$ is a purifying system. We can formulate $2^{-H_{\min}^{\tilde{\varepsilon}+\tilde{\varepsilon}'}}(A'|B')_{\bar{\rho}}$ as the following semidefinite program:

PRIMAL PROBLEM:

$$\begin{aligned} \text{minimum:} \quad & \operatorname{tr}[\mathbb{I}_{B'}\,\sigma_{B'}] \\ \text{subject to:} \quad & \mathbb{I}_{A'}\otimes\sigma_{B'} \geq \operatorname{tr}_C[\tilde{\rho}_{A'B'C}] \\ & \operatorname{tr}[\tilde{\rho}_{A'B'C}\bar{\rho}_{A'B'C}] \geq 1-(\tilde{\varepsilon}+\tilde{\varepsilon}')^2 \\ & \operatorname{tr}[\tilde{\rho}_{A'B'C}] \leq 1 \\ & \sigma_{B'} \geq 0,\ \tilde{\rho}_{A'B'C} \geq 0 \end{aligned}$$

DUAL PROBLEM:

$$\begin{aligned} \text{maximum:} \quad & (1-(\tilde{\varepsilon}+\tilde{\varepsilon}')^2)\lambda - \mu \\ \text{subject to:} \quad & \operatorname{tr}_A[E_{A'B'}] \leq \mathbb{I}'_B \\ & \lambda\bar{\rho}_{A'B'C} \leq E_{A'B'}\otimes\mathbb{I}_C + \mu\,\mathbb{I}_{A'B'C} \\ & E_{A'B'} \geq 0,\ \lambda,\ \mu \geq 0, \end{aligned}$$

where $\sigma_{B'}$ and $\tilde{\rho}_{A'B'C}$ are the primal variables and $E_{A'B'}$, λ and μ are the dual variables, respectively. Let $Z_{A'B'}$ be a primal optimal plan for the semidefinite program of $H_{\max}(A'|B')_{\bar{\rho}}$, that is $Z_{A'B'}\otimes\mathbb{I}_C \geq \bar{\rho}_{A'B'C}$ and $\operatorname{tr}_{A'}[Z_{A'B'}] \leq 2^{H_{\max}(A'|B')_{\bar{\rho}}}\mathbb{I}_{B'}$. Then the variables $E_{A'B'} = 2^{-H_{\max}(A'|B')_{\bar{\rho}}}Z_{A'B'}$, $\lambda = 2^{-H_{\max}(A'|B')_{\bar{\rho}}}$ and $\mu = 0$ are a dual feasible plan for the above semidefinite program. By the weak duality theorem we have then

$$(1-(\tilde{\varepsilon}+\tilde{\varepsilon}')^2)2^{-H_{\max}(A'|B')_{\bar{\rho}}} \leq 2^{-H_{\min}^{\tilde{\varepsilon}+\tilde{\varepsilon}'}(A'|B')_{\bar{\rho}}}.$$

Taking the logarithm and considering the fact that all states which are $\tilde{\varepsilon}'$-close to $\hat{\rho}_{A'B'}$ are contained in the $(\tilde{\varepsilon}+\tilde{\varepsilon}')$-neighborhood of $\bar{\rho}_{A'B'}$, we get

$$H_{\min}^{\tilde{\varepsilon}'}(A'|B')_{\hat{\rho}} \leq H_{\min}^{\tilde{\varepsilon}+\tilde{\varepsilon}'}(A'|B')_{\bar{\rho}} \leq H_{\max}^{\tilde{\varepsilon}}(A'|B')_{\hat{\rho}} + \log\left(\frac{1}{1-(\tilde{\varepsilon}+\tilde{\varepsilon}')^2}\right). \tag{30}$$

By Proposition 5.3 in [16] we have $H_{\min}^{\tilde{\varepsilon}'}(A'|B')_{\hat{\rho}} = H_{\min}^{\tilde{\varepsilon}'}(A|B)_{\hat{\rho}}$ and $H_{\max}^{\tilde{\varepsilon}}(A'|B')_{\hat{\rho}} = H_{\max}^{\tilde{\varepsilon}}(A|B)_{\hat{\rho}}$. Finally, substituting in (30) $\tilde{\varepsilon} = \varepsilon + \sqrt{1-\operatorname{tr}(\rho_{AB})}$ and $\tilde{\varepsilon}' = \varepsilon' + \sqrt{1-\operatorname{tr}(\rho_{AB})}$ and considering that $H_{\min}^{\varepsilon'}(A|B)_\rho \leq H_{\min}^{\varepsilon'+\sqrt{1-\operatorname{tr}\rho}}(A|B)_{\hat{\rho}}$ as well as $H_{\max}^{\varepsilon+\sqrt{1-\operatorname{tr}\rho}}(A|B)_{\hat{\rho}} \leq H_{\max}^{\varepsilon}(A|B)_\rho$ we conclude the proof. ∎

Appendix B
Technical Lemmas

A. Operator inequalities

Theorem 17 ([17], Theorem 1)**.** *Let Q and R be positive operators on a Hilbert space $\mathcal{H}$ and let $0 \leq s \leq 1$. Then,*

$$\operatorname{tr}\left[Q^s R^{1-s}\right] \geq \frac{1}{2}\operatorname{tr}\left[Q+R-|Q-R|\right] \tag{31}$$

From this theorem we can draw the following useful corollary.

Corollary 18. *Let R and Q be positive operators on a Hilbert space $\mathcal{H}$, let $0 \leq s \leq 1$ and let $P_{\pm}$ denote the orthogonal projectors onto the eigenspaces corresponding to positive/negative eigenvalues of the operator $Q - R$, respectively. Then,*

$$\operatorname{tr}\left[Q^s R^{1-s}\right] \geq \operatorname{tr}\left[P_+ R + P_- Q\right]$$

Proof: We make the following decomposition of $|Q - R|$

$$|Q - R| = P_+ (Q - R) P_+ - P_- (Q - R) P_-, \tag{32}$$

where $P_{\pm}$ are the projectors onto the positive and negative eigenvalues of $Q - R$, respectively. Substituting (32) in (31) and using the fact that $P_+ + P_- = \mathbb{I}$, we obtain

$$\begin{aligned}\operatorname{tr}\left[Q^s R^{1-s}\right] &\geq \frac{1}{2}\operatorname{tr}\left[Q + R - |Q - R|\right] \\ &= \operatorname{tr}\left[P_- Q + (\mathbb{I} - P_-) R\right] \\ &= \operatorname{tr}\left[P_- Q + P_+ R\right].\end{aligned}$$

■

B. Purified Distance: Properties

Lemma 19 ([7], Lemma 7)**.** *If $\rho, \sigma \in \mathcal{S}_{\leq}(\mathcal{H})$ and $\mathcal{E}$ is a trace non-increasing CPM on $\mathcal{L}(\mathcal{H})$, then*

$$P(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \leq P(\rho, \sigma).$$

Evidently, for any $0 \leq \Pi \leq 1$ the map defined by $\rho \longmapsto \Pi\rho\Pi$, $\rho \in \mathcal{S}_{\leq}(\mathcal{H})$ is a trace non-increasing CPM. Thus, in particular, by the above lemma we have

$$P(\Pi\rho\Pi, \Pi\sigma\Pi) \leq P(\rho, \sigma) \tag{33}$$

for ρ, $\sigma \in \mathcal{S}_{\leq}(\mathcal{H})$.

Lemma 20 ([18], Lemma 7)**.** *Let $\rho \in \mathcal{S}_{\leq}(\mathcal{H})$ and $0 \leq \Pi \leq \mathbb{I}$. Then,*

$$P(\Pi\rho\Pi, \rho) \leq \frac{1}{\sqrt{\operatorname{tr}\rho}} \sqrt{(\operatorname{tr}\rho)^2 - (\operatorname{tr}[\Pi^2\rho])^2}.$$

When Π is a projector, that is $\Pi^2 = \Pi$, then a straightforward computation yields

$$P(\Pi\rho\Pi, \rho) \leq \sqrt{2\operatorname{tr}[\Pi^{\perp}\rho] - (\operatorname{tr}[\Pi^{\perp}\rho])^2} \tag{34}$$

where $\Pi^{\perp} = \mathbb{I} - \Pi$ is the orthogonal complement of Π.

Lemma 21 ([7], Lemma 8)**.** *Let $\rho, \sigma \in \mathcal{S}_{\leq}(\mathcal{H})$, $\mathcal{H}' \cong \mathcal{H}$ and $\bar{\rho} \in \mathcal{S}_{\leq}(\mathcal{H} \otimes \mathcal{H}')$ be a purification of ρ. Then, there exists a purification $\bar{\sigma} \in \mathcal{S}_{\leq}(\mathcal{H} \otimes \mathcal{H}')$ of σ such that $P(\bar{\rho}, \bar{\sigma}) = P(\rho, \sigma)$.*

From that lemma one infers the following corollary:

Corollary 22. *Let $\rho, \sigma \in \mathcal{S}_{\leq}(\mathcal{H})$, $\mathcal{H}' \cong \mathcal{H}$ and $\bar{\rho} \in \mathcal{S}_{\leq}(\mathcal{H} \otimes \mathcal{H}')$ be an extension of ρ. Then, there exists an extension $\bar{\sigma} \in \mathcal{S}_{\leq}(\mathcal{H} \otimes \mathcal{H}')$ of σ such that $P(\bar{\rho}, \bar{\sigma}) = P(\rho, \sigma)$.*

REFERENCES

[1] C. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, 1948.

[2] R. Renner and S. Wolf, "Smooth Rényi entropy and applications," *Proc. IEEE Int. Symp. Info. Theory*, p. 233, 2004.

[3] R. Renner, "Security of quantum key distribution," Ph.D. dissertation, ETH Zürich, 2005. [Online]. Available: http://arxiv.org/abs/quant-ph/0512258v2

[4] R. Renner and R. König, "Universally composable privacy amplification against quantum adversaries," *Springer Lecture Notes in Computer Science*, vol. 3378, no. 9, pp. 407–425, 2005.

[5] R. König, R. Renner, and C. Schaffner, "The operational meaning of min- and max-entropy," *IEEE Transactions on Information Theory*, vol. 55, no. 9, pp. 4674–4681, 2009.

[6] M. Tomamichel, R. Colbeck, and R. Renner, "A fully quantum asymptotic equipartition property," *IEEE Transactions on Information Theory*, vol. 55, pp. 5840–5847, 2009.

[7] ——, "Duality between smooth min- and max-entropies," *IEEE Transactions on Information Theory*, vol. 56, pp. 4674–4681, 2010.

[8] N. Datta, "Min- and max-relative entropies and a new entanglement monotone," *IEEE Transactions on Information Theory*, vol. 55, no. 6, p. 2816, 2009.

[9] L. del Rio, J. Aberg, R. Renner, O. C. O. Dahlsten, and V. Vedral, "The thermodynamic meaning of negative entropy," *Nature*, vol. 474, no. 7349, pp. 61–63, 2011.

[10] F. Dupuis, "The decoupling approach to quantum information theory," Ph.D. dissertation, Université de Montréal, Apr. 2009. [Online]. Available: http://arxiv.org/abs/1004.1641

[11] F. Dupuis, M. Berta, J. Wullschleger, and R. Renner, "The decoupling theorem," Dec. 2010. [Online]. Available: http://arxiv.org/abs/1012.6044

[12] M. Berta, "Single-shot quantum state merging," Master's Thesis, ETH Zurich, 2008.

[13] A. Barvinok, *A Course in Convexity*, ser. Graduate Studies in Mathematics. American Mathematical Sociaty, 2002, vol. 54.

[14] J. Watrous, "Theory of quantum information," Fall 2011, lecture notes. [Online]. Available: http://www.cs.uwaterloo.ca/~watrous/CS766/

[15] A. Uhlmann, "The transition probability in the state space of a *-algebra," *Rep. Math. Phys.*, vol. 9, no. 273, 1976.

[16] M. Tomamichel, "A framework for non-asymptotic quantum information theory," Ph.D. dissertation, ETH Zürich, 2012. [Online]. Available: http://arxiv.org/abs/arXiv:1203.2142

[17] K. Audenaert, J. Calsamiglia, R. Muñoz-Tapia, E. Bagen, L. Masanes, A. Acin, and F. Verstraete, "Discriminating states: The quantum Chernoff bound," *Physical Letters Review*, vol. 98, pp. 160 501–4, 2007.

[18] M. Berta, M. Christandl, R. Colbeck, J. Rennes, and R. Renner, "The uncertainty principle in the presence of quantum memory," *Nature Physics*, vol. 1734, 2010.